虎尾科技大學
40723115
林于哲

April 1 2021

目錄

Introduction

- Activation Function
  I. Sigmoid Function
  II. Relu
- Optimizer
  I. Adam
  II. SigmoidPrime
- Loss function
  I. Mean Squared Error

1 What is Reinforcement Learning?

強化學習是通過 agent 與已知或未知的環境持互動，不斷適應與學習，得到的回饋可能是正面，也就是 reward，如果得到負面，那就是 punishments。考慮到 agent 與環境互動，我們就能決定要執行哪個動作。簡而言之，強化學習是建立在 reward 與 punishments 上。

The key point of Reinforcement Learning:

- It differs from normal Machine Learning, as we do not look at training datasets.
- Interaction happens not with data but with environments, through which we depict real-world scenarios
- As Reinforcement Learning is based on environments, many parameters come in to play. It takes lots of information to learn and act accordingly.
- Environments in Reinforcement Learning are real-world scenarios that might be 2D or 3D simulated worlds or gamebased scenarios.
- Reinforcement Learning is broader in a sense because the environments can be large in scale and there might be a lot of factors associated with them.
- The objective of Reinforcement Learning is to reach a goal.
- Rewards in Reinforcement Learning are obtained from the environment.

2 Faces of Reinforcement Learning

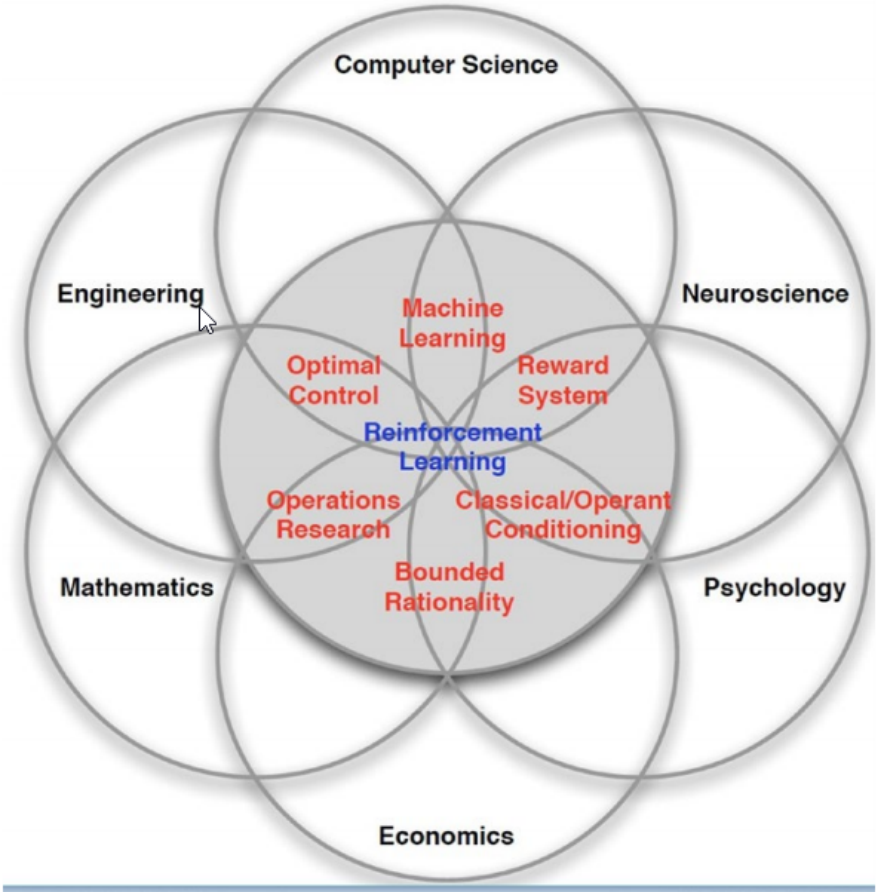

fig 1: Venn diagram;

3 The Flow of Reinforcement Learning

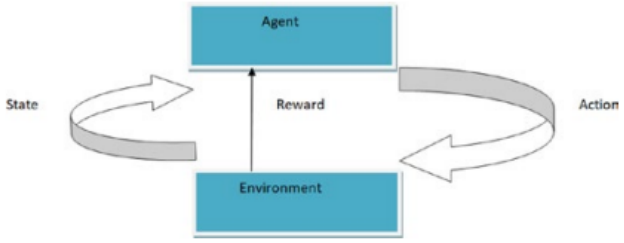

fig 2: RL structur;

The key points of consideration:

- The Reinforcement Learning cycle works in an interconnected.
- The distinct communication happens with rewards in mind.
- There is distinct communication between the agent and the environment.
- The object or robot moves from one state to another.
- An action is taken to move from one state to another

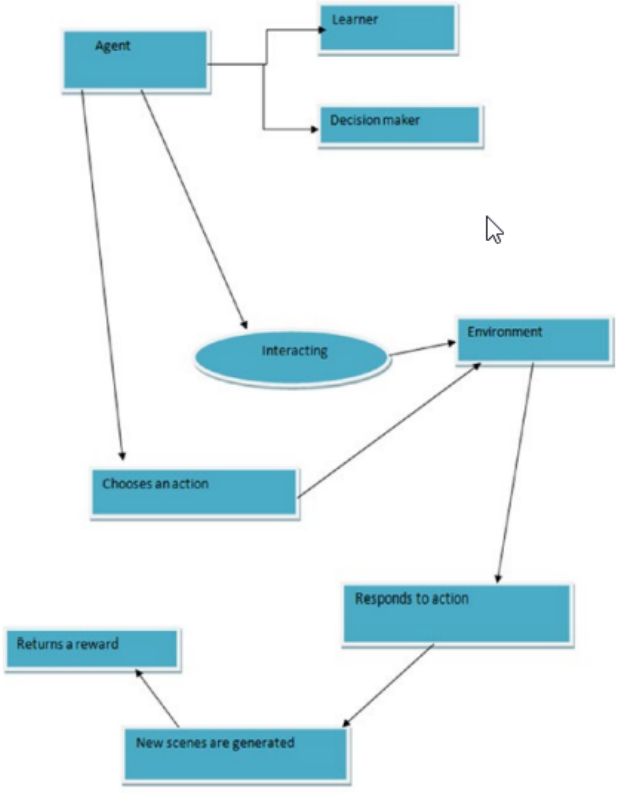

fig 3: The entire interaction process;

The agent is also a decision maker because it tries to take an action that will get it the maximum reward.
When the agent starts interacting with the environment, it can choose an action and respond accordingly. From then on, new scenes are created. When the agent changes from one place to another in an environment, every change results in some kind of modification. These changes are depicted as scenes. The transition that happens in each step helps the agent solve the Reinforcement Learning problem more effectively

fig 4: The entire interaction process;

4 Different Terms in Reinforcement Learning

There are two constants that are important in this case—gamma (γ) and lambda (λ)

Gamma is used in each state transition and is a constant value at each state change. Gamma allows you to give information about the type of reward you will be getting in every state. Gamma (γ) is called a discount factor and it determines what future reward types we get:

- A gamma value of 0 means the reward is associated with the current state only.
- A gamma value of 1 means that the reward is long-term.

Lambda is generally used when we are dealing with temporal difference problems. It is more involved with predictions in successive states. Increasing values of lambda in each state shows that our algorithm is learning fast. The faster algorithm yields better results when using Reinforcement Learning techniques. As you' ll learn later, temporal differences can be generalized to what we call TD(Lambda).

5 Interactions with Reinforcement Learning

The interactions between the agent and the environment occur with a reward. We need to take an action to move from one state to another.
Reinforcement Learning is a way of implementing how to map situations to actions so as to maximize and find a way to get the highest rewards. The machine or robot is not told which actions to take, as with other forms of Machine Learning, but instead the machine must discover which actions yield the maximum reward by trying them.